

Automating the process of identifying the preferred representational system in Neuro Linguistic Programming using Natural Language Processing

Main Author and Correspondent Author:

Mohammad Hossein Amirhosseini

**PhD researcher, MSc, BSc, School of Computing and Digital Media, London Metropolitan
University, Tower Building, 166-220 Holloway Road, London N7 8DB, UK.**

E-mail: m.amirhosseini@londonmet.ac.uk

Professor Hassan Kazemian

**PhD, FBCS, FIET. School of Computing and Digital Media, London Metropolitan University, Tower
Building, 166-220 Holloway Road, London N7 8DB, UK.**

E-mail: h.kazemian@londonmet.ac.uk

Abstract

Neuro Linguistic Programming (NLP) is a methodology used for recognition of human behavioural patterns and modification of the behaviour. A significant part of this process is influenced by the theory of representational systems which equates to the five main senses. The preferred representational system of an individual can explain a large part of exhibited behaviours and characteristics. There are different methods to recognise the representational systems, one of which is to investigate the sensory based words in the used language during the conversation. However, there are difficulties during this process since there is not a single reference method used for identification of representational systems and existing ones are subject to human interpretations. Some human errors like lack of experience, personal judgment, different levels of skill and personal mistakes may also affect the accuracy and reliability of the existing methods. This research aims to apply a new approach that is to automate the identification process in order to remove human errors thereby increasing the accuracy and precision. Natural Language Processing has been used for automating this process and an intelligent software has been developed able to identify the preferred representational system with increased accuracy and reliability. This software has been tested and compared to human identification of representational systems. The results of the software are similar to a NLP practitioner and the software responds more accurately than a human practitioner in various parts of the process. This novel methodology will assist the NLP practitioners to obtain an improved understanding of their clients' behavioural patterns and the associated cognitive and emotional processes.

Key words: Neuro Linguistic Programming, Natural Language Processing, representational systems, behavioural patterns, communication improvement, text processing

1. Introduction

Neuro Linguistic programming (NLP) is recognised as a collection of techniques that can help to identify the way in which people think, how they use words and language to communicate and behave moreover, detecting patterns in people's behaviour (Transform Destiny 2008). On this basis, NLP provides a series of techniques, communication tools, approaches and attitudes that can help people and organisations achieve their goals in management and personal development (Lazarus 2010). Another definition by Casale (2012) defined NLP as 'A school of psychological techniques that effectively communicate with the listener's subconscious or unconscious mind'.

NLP has been deployed by well-known companies such as NASA, IBM, McDonald's and the U.S. Army (Witkowski 2010) and it has also been mentioned that NLP is being applied widely and often informally in UK education (Singer and Lalich 1996). NLP has become popular amongst the majority of psychologists and also university employees (Tosey and Mathison 2003). Thus, the application of NLP has had success across different disciplines, thereby increasing confidence in its utility.

NLP was developed in the early 1970s by Bandler and Grinder. Initially NLP focused on the strategy people used to process information and how this strategy can be recognised and understood. This then developed into a collection of tools, techniques and frameworks to be used in different disciplines (Tosey and Mathison 2006). There are variety of techniques included in NLP with varying steps in the personal development process. However, one of the most important stages is the identification of the preferred representational system of an individual. In the context of NLP, representational systems are the different ways that we represent or store information in our mind (Ellerton 2007). This occurs via the five main sensory modalities through which people comprehend the world, coding and storing information in their mind through seeing, hearing,

feeling, tasting and smelling and then filter this information with their beliefs and values in order to re-represent experience to themselves and finally act on the result (O'Connor and Seymour 1993; Linder-Pelz 2010). Therefore, through examining the representational systems in NLP, we can assess how the human mind processes information and interprets meanings (McAfee 2014). Palmiero, Di Matteo and Benardinelli (2014) also point out that how people represent conceptual knowledge is a long-debated issue and one important approaches assumes that conceptual knowledge is distributed across different attribute domains, such as vision, touch.

While we use all sensory based representational systems as a means for learning, each one of us has got a dominant preferred system that we use more than others. This preferred system is conveyed through different ways in an individual's speech, learning methods, and other communicatory pathways (NLP Dynamics Ltd. 2013). There are different generalizations of characteristics which can be related to people with preference to each of the representational systems. Hence, understanding the preferred representational system of an individual reveals a lot about likely characteristics, behavioural patterns and learning processes which can be key to NLP modelling and personal development processes (NLP Dynamics Ltd 2013).

Ellerton (2007) suggests that there are six representational systems in total instead of five. Five of them correspond to the main senses which are visual, auditory, kinesthetic, olfactory and gustatory (VAKOG) and the sixth one is identified as the auditory digital representational system which is a non-sensory system and refers to how people sort experience following its occurrence (Monkeypuzzle training and consultancy 2016) focusing on self-talk, discrete words, facts, figures and logic. Ellerton (2015) also proposes that people often work with three representational systems, the visual, auditory and kinesthetic (VAK) and two other representational systems - gustatory and olfactory, do not play a major role and are often included within kinesthetic. Through

a NLP study, however, Rayner Institute (2015) recognises four primary representational systems in total. Smell and taste were disregarded as they are not normally a primary sense for most people, and the category of auditory digital was added. This study led NLP practitioners to confine their consideration to the VAK representational systems and the digital representational system when assessing for the preferred representational system.

Each representational system is associated with specific tendencies of characteristics. McAfee (2014) explains that visual people usually memorise via observation of imagery and they are interested to see how a concept looks like. They are less distracted by noise and have trouble remembering long verbal instructions. In other words, what they see has a priority and is more important than what they understand and experience through hearing or feeling (Monkeypuzzle training and consultancy 2016). On the other hand, Bensted (2014) discusses that auditory people typically are easily distracted by noise. They can learn and memorize by listening and tone of voice can be very important to them. They like music and can repeat things easily. For this group, what they hear has a priority and is more important than what they understand or experience through seeing or feeling (Monkeypuzzle training and consultancy 2016).

People with a kinesthetic preference memorise by doing or walking through something. They are more interested in a program that gives them a gut feeling or in something that feels right (McAfee 2014). They also respond very well to physical rewards and touching (Bensted 2014). This group loves physical activities and they are more interested in trying something out and less interested in theory (Monkeypuzzle training and consultancy 2016).

Finally, people with an auditory digital preference, spend a fair amount of time talking to themselves (Bensted. 2014). They usually memorize by steps, procedures and sequences and it is really important for them to know if the program makes sense (McAfee 2014). In fact, for this

group Logic is a priority and is more important than how they understand and experience through seeing, hearing or feeling (Monkeypuzzle training and consultancy 2016). In other words, they are more interested in facts and science and before doing something, they need to understand it (Monkeypuzzle training and consultancy 2016).

Language is recognised as a key identifier of dependence to sensory modalities. Recognition of sensory words termed ‘predicates’ in language can reveal the use of the related sensory modality and thus give an indication of an individual’s preferred system of use. Accordingly, adapting the language used to match an individual’s, based on their preferred system, will assist them in understanding what you wish to communicate (Brefi Group Limited 2004). Thus, the preferred representational system can be recognised through analysis of the language used in conversation, considering the sentences and words used by an individual for predicates. There are defined patterns for identification of the preferred representational system by psychologists and NLP practitioners. This method, however, is unguarded against human factors such as lack of experience, personal judgment, mistakes and inaccuracy that will have a direct or indirect impact on the identification of systems.

There have been previous attempts to automate the identification process of the preferred representational systems. However, they have only been as competent as a simple computerization of the concept rather than intelligent automation. These methods are akin online self-assessment questionnaires and where answers are based on the individual’s judgment and opinion of themselves or often already provide discrete options to be chosen from rather than allowing a more candid expression. Moreover, some services, although classified as ‘automatic’, do not provide immediate results, sometimes requiring answers to be sent to a NLP practitioner for analysis and the results then being sent to clients. Another shortcoming of the available online surveys is their

simplicity, which results in limited considerations and ultimately, reduced accuracy or error. Additionally, none of the services available use artificial intelligence in their attempts of automation. During the literature review of this research, no windows application or online application was found with the ability of having a conversation with a human being, or in other words, the ability of computer-human interaction.

Intelligent systems are created in order to work well, in different situations and environments and their intelligence enables them reach the maximum probability of success even if there is not enough knowledge of the situation. As a result, artificial intelligence can be considered as a powerful tool for automating a process in this situation and where a human needs to interact with a computer directly (Gudwin 2000). According to Chopra, Prashar and Sain (2013) Natural Language Processing is a subfield of artificial intelligence and linguistic which try to enables computers understand the words or sentences written in human languages. In fact, Natural Language Processing is a field of study which is related to different areas of research like computer science, artificial intelligence, linguistics and psychology and it mainly focuses on the interactions between computers and human languages (Chopra, Prashar and Sain 2013). It allows exploration of how computers can be used to comprehend and manipulate the natural language of speech and text to reach outcomes (Chowdhury 2003). Natural Language Processing may begin at the word-level to determine the structure and nature (such as part-of-speech or meaning) of the word. It can, however, be developed for identification of meaning in sentences whereby word order, grammar and meaning of the entire sentence can be recognised. Moreover, it is possible to further develop a Natural Language Processing system to acknowledge the context and the overall environment or domain (Chowdhury 2003). On this basis, Natural Language Processing was recognised as the most competent tool for identification of the predicates and sensory based language in the human-

computer communication for detection of the preferred representational system.

According to Dale (2010) there are 5 stages of analysis in processing the natural language which are tokenization, lexical analysis, syntactic analysis, semantic analysis and pragmatic analysis after which the speaker's intended meaning can be understood. There are various Natural Language Processing tools that can be useful in these processes. OpenNLP is a set of tools which are created based on machine learning techniques and it can aid the most common Natural Language Processing tasks such as sentence detection, tokenization, Part of Speech or POS tagging and other relevant tasks (Hervás et al. 2012). WordNet is another tool including a vast lexical database in English and in this database, nouns, verbs, adjectives and adverbs are grouped into sets of synonyms (Hervás et al. 2012).

This research attempts to develop a new methodology for a more comprehensive, competent process of detecting the preferred representational system. This is carried out by attempting to remove the limitations found in the manual and computerized services available. In this light, human factors such as lack of skills and experience, personal judgment, inaccuracy, human error and similar factors associated with analysis by psychologists and NLP practitioners are eliminated. Alternatively, an intelligent system with the ability to analyze natural language is developed with the capability of acknowledging the meaning of the words, sentences and context used in order to detect the pattern of the language associated with the preferred representational system as a more accurate and reliable NLP method.

2. Methodology

2.1 Application structure and development procedure

As discussed before, Natural Language Processing has been perceived as the most relevant and powerful technique for automating the preferred representational system identification process. The software was developed using Python which is a very flexible programming language as well as NLTK which is a very powerful Natural Language Toolkit for Python. According to the NLTK official website, NLTK is a leading toolkit for developing Python programs utilizes human language data. NLTK has a user friendly interfaces for over 50 corpora and lexical resources like WordNet. NLTK has a suite of text processing libraries for classification, tokenization, stemming, tagging, parsing, and semantic reasoning, wrappers for industrial-strength NLP libraries, and an active discussion forum (NLTK official website 2016).

The software is able to have a conversation with human beings and communicate through an interactive environment starting with a brief introduction followed by response-based questions. The individual's answers are communicated through typing out of relevant response. The answers will be analyzed by the software within and in the end of the conversation. Based on an overall analysis, the software will then recognise the preferred representational system and based on the relevant characteristics to that representational system, suggest relevant solutions for improvement in communication and learning.

2.1.1 Tokenization process

The first step in analysis of response is via the Tokenization technique. According to Manning, Raghavan and Schütze (2009) 'a token is an instance of a sequence of characters in some particular document that are grouped together as a useful semantic unit for processing'. In fact, these basic

units must be clearly segregated, otherwise it will be impossible to carry out any analysis or generation (Webster and Kit 1992).

Hence, the individual's answer will be recorded as a string and this string will be divided to different sentences and each sentence analyzed separately. All these sentences will be recorded in a list and each sentence will be divided by words and all these words will be recorded in a different list again. These lists will be used for lexical and syntactic analysis in the next step.

2.1.2 Lexical and Syntactic analysis

A one-dimensional language like a written language as composed of letters and symbols, can be considered as a code describing some reality (Horn, 2008). It needs to have rules in order to describe how its words or sentences are connected to that reality and how to put them together into a language representation (Horn, 2008). As a result, Part-Of-Speech tagging or POS tagging technique will be used in this step. In this process, the software will recognise the role of each word in each sentence. Jurafsky and Martin (2014) explained that in this process a part-of-tagging speech marker will be assigned to each word in an input text. Bird, Klein and Loper (2009) also in a book named 'Natural Language Processing by Python' stated that a part-of-speech tagger can process a sequence of words and then attach a part of speech tag to each one of those words. There are different lists of Parts-Of-Speech tag sets and one of the most common tag sets is Brown corpus which was used in this research. As a result, all verbs, nouns, adjectives, adverbs and other relevant elements in each sentence will be recognised.

Hidden Markov model (HMM) which is a very common and useful tagging technique was used in this research to build the POS tagger. Below is the bigram HMM equation:

$$x_m = \arg \max_n P(x_n | x_{m-1}, y_m) \quad (1)$$

The nearby words and tags will be checked in order to solve the tagging problem:

$$x_m = \arg \max_n P(x_n | x_{n-1})P(y_m | x_n) \quad (2)$$

In this equation $P(y_m | x_n)$ is for word likelihood and $P(x_n | x_{n-1})$ is for tag co-occurrence. The full model will try to identify the best sequence of tags for the whole sentence:

$$\begin{aligned} \hat{X} &= \arg \max_{X \in \tau} P(X | Y) \\ &= \arg \max_{X \in \tau} \frac{P(X)P(Y | X)}{P(Y)} \\ &= \arg \max P(X)P(Y | X) \end{aligned} \quad (3)$$

The chain rule will be used to expand this equation:

$$\begin{aligned} P(X)P(Y | X) &= \\ \prod_{m=1}^k P(y_m | y_1 x_1 \dots y_{m-1} x_{m-1} x_m) P(x_m | y_1 x_1 \dots y_{m-1} x_{m-1}) \end{aligned} \quad (4)$$

In order to approximate these two factors, the trigram assumption will be simplified. The probability of a word depends only on its tag:

$$P(x_m | y_1 x_1 \dots x_{m-1} x_m) = P(y_m | x_m) \quad (5)$$

Then tag history will be approximated by the two most recent tags:

$$P(x_m | y_1 x_1 \dots x_{m-1}) = P(x_m | x_{m-2} x_{m-1}) \quad (6)$$

Eventually, the equation will be replaced:

$$\begin{aligned} P(X)P(Y | X) &= \\ P(x_1)P(x_2 | x_1) \prod_{m=3}^k P(x_m | x_{m-2} x_{m-1}) & \left[\prod_{m=1}^k P(y_m | x_m) \right] \end{aligned} \quad (7)$$

Following the process of POS tagging, steaming technique will be used to remove all prefixes

and suffixes from the words to identify the root of each word. Manning, Raghavan and Schütze (2009) explained that because of grammatical reasons, each document may use different forms of a word and also there are families of derivationally related words that may have similar meanings. Hence, in different situations, it might be useful if we can search for one of these words as a root word. For instance, it may help to return some other documents that contain another word related to the root word (Manning, Raghavan and Schütze 2009). Accordingly, by using this technique, all the roots of the words will be recognised and recorded in a new list for the comparison process in the next step.

2.1.3 Comparison process

There are six other lists that have been defined for the software, whereby each one of these lists are allocated to one specific representational system, including a collection of words which are the associated predicates for the specified system. A collection of various documents has been used in order to make a collection of relevant vocabulary. These include ‘NLP Home Study Programme (V2.0)’ published by Juiced Concepts Limited (2012), ‘Representational Systems’ published by Brefi Group Limited (2004) and ‘The Power of Words’ written by Katy McAfee (2014). Tables 1-4 below are showing some of the most common predicates recorded in the relevant lists.

Table 1 Example of visual representational system predicates

Visual predicates list			
See	Saw	Seen	Look
Appear	Observing	Appearance	View
Show	Shown	Dawn	Reveal
Envision	Illuminate	Twinkle	Clear
Foggy	Focus	Hazy	Crystal
Flash	Image	Picture	Spark
Frame	Snapshot	Vivid	Imagine
Clarity	Witness	Illustrate	Vague
Outlook	Inspect	Sight	Light
Scene	Watch	Perceive	Perspective
Observe	Vision	Angle	Sign
...

Table 2 Example of auditory representational system predicates

Auditory predicates list			
Hear	Listen	Sound	Music
Harmonize	Tune in	Tune out	Ear
Ring	Bell	Silence	Heard
Resonate	Deaf	Mellifluous	Dissonance
Dissonant	Overtones	Attune	Outspoken
Tell	Announce	Remark	Overtones
Unhearing	Audible	Voice	Interview
Talk	Speak	Rumor	State
Whine	Babble	Echo	Orchestrate
Whisper	Oral	Hum	Speechless
...

Table 3 Example of kinesthetic representational system predicates

Kinesthetic predicates list			
Feel	Touch	Grasp	Catch
Hold	Contact	Throw out	Hard
Feeling	Concrete	Scrape	Handle
Suffer	Impression	Flow	Lukewarm
Slip	Tap	Shift	Throw
Turn around	Unfeeling	Callous	Solid
Unjudging	Softly	Soft	Rub
Unsettles	Smooth	Pushy	Push
Panicky	Stumble	Muddled	Relaxed
Relax	Loose	Sore	Bearable
Cool	Tepid	Charge	Heavy
...

Table 4 Example of auditory digital representational system predicates

Auditory digital predicates list			
Sense	Experience	Understand	Catch
Learn	Process	Decide	Hard
Consider	Change	Perceive	Handle
Distinct	Conceive	Know	Lukewarm
Conscious	Recall	Communicate	Throw
Plan	Advice	Function	Solid
Create	Activate	Repeat	Rub
Logically	Reasonable	Statistically	Push
...

NLTK or Natural Language Toolkit in Python was the method used to help in identifying all possible synonyms for each one of these representational system predicates and record them in a separate list. As a result, we acquired a comprehensive collection of all possible representational

system predicates. Thus, in this step the software would be able to compare the obtained list, explained in section 2.1.2, containing the roots of word, with the 6 lists associated with each representational system and their synonyms. If any of words in the list being analyzed was in any of representational system predicates lists or their synonym lists, the counter for that specific representational system will be increased. Finally, all counters will be compared together and the application will inform the person about his or her preferred representational system. Moreover, some solutions will be suggested to improve the communication effectiveness and learning methodologies of the individual. The relevant solutions have been collected from the aforementioned documents. Figure 1 shows the software development procedure.

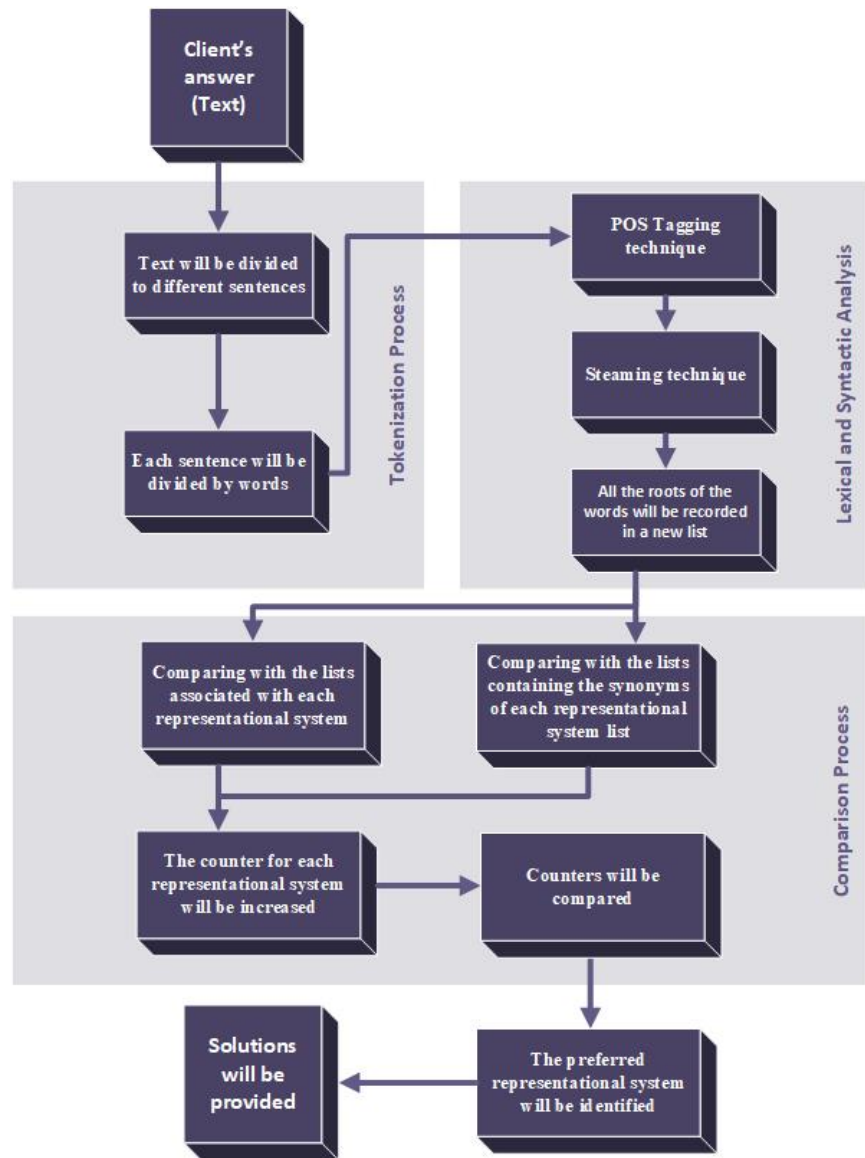


Fig. 1 Software development procedure

2.2 Participants and Data gathering procedure

This intelligent software has been tested on a group of 55 students at London Metropolitan University in different levels of study including 14 PhD researchers, 17 Masters students and 24

Bachelors students. Some of these students had additional work experience in industry. 31 students were female and 24 students were male, their ages ranging from 18 to 34. 23 participants were between 18 to 21 which is 41.8 percent of all participants. 14 people were between 22 to 25, 12 people were between 26 to 29 and finally 6 participants were between 30 to 34 years old. So majority of participants were in the first group which is 18 to 21.

The software has also been tested in the UKSTUDY Company, located in Brighton. This company consists of 14 employees based in the office and 10 other temporary project based employees. All 24 employees contributed to the data gathering process.

Before interaction with the software, participants were fully informed about this software and its application. They were also convinced that all the information provided will stay private and confidential and not shared with any third parties. Also that their answers will not be used for any other purpose than analysis of the accuracy and performance of the software. They were also aware that there are no affects such as financial loss, mental or physical harm as a consequence of their participation and their answers will not have any effect on their study, their job or personal life.

For the group of 55 students at London Metropolitan University, the time needed for having a conversation with the software ranged between 10 to 15 minutes depending on the length of their answers and their typing speed. For the employees at the UKSTUDY Company, the duration of the conversation with the software was between 12 to 16 minutes.

They were also asked to respond to a multiple choice questionnaire for representational system preference, designed by Steve Antcliff in a book named 'Life coaching-made simple' published in 2009. This questionnaire contains 12 multiple-choice questions providing 4 possible answers. Each answer indicates preference of a specific representational system but the order of

answers is different for each question. For instance, in question 1, the first option is associated with the kinesthetic, the second to the auditory, the third to the auditory digital and the fourth answer to the visual representational system. In question 2, however, the order of association of the 4 answers to representational systems is the visual, the auditory, the kinesthetic and the auditory visual representational system respectively. Participants are asked to rate these options by choosing a number from 1 (As the least descriptive of them) to 4 (As the closest answer to describe them).

This questionnaire has an answer sheet to calculate the final score for each representational system and the preferred representational system can be identified based on the highest score between the categories. Figure 2 shows the answer sheet for this questionnaire.

Question Number	Visual	Auditory	Kinaesthetic	Auditory Digital
1	d	b	a	c
2	a	b	c	d
3	a	d	b	c
4	b	d	a	c
5	c	d	b	a
6	c	a	d	b
7	c	a	b	d
8	d	b	c	a
9	b	c	d	a
10	a	c	d	b
11	d	a	c	b
12	b	c	a	d
total	V =	A =	K =	A _d =

Fig. 2 Answer sheet for the preferred representational system questionnaire

This questionnaire was chosen after considering four other popular representational system questionnaires offered in various documents including ‘NLP Home Study Programme (V2.0)’ published by Juiced Concepts Limited (2012), ‘Representational Systems’ published by Brefi Group Limited (2004), ‘Introduction to Neuro-Linguistic Programming’ published by Transformed Destiny (2015) and ‘The Power of Words’ written by Katy McAfee (2014). It was realized that these questionnaires are a simplified version of the chosen questionnaire with less number of questions with some not considering the auditory digital representational system. The results of the chosen questionnaire was compared with the results of our intelligent software for similarities and examination of the software’s performance and accuracy. The software performance was also compared with some of the previous automation attempts.

2.3 Data analysis procedure

After the data gathering process, the accuracy and the performance of the software were evaluated. Answers for each individual was recorded in a separate file and all representational system predicates were extracted from the text manually. These extracted words were then compared with the extracted words by the software that have been achieved automatically. Finally, the results were calculated and these results were presented in figures shown in the result section. Moreover, the preferred representational system identified by the software were compared to the results of the manual questionnaire for similarity to analyse the accuracy. The performance of the software has also been compared with the previous automation attempts.

3. Results and discussion

3.1 Comparing the results of software and questionnaire for the academic group

As it was mentioned in data gathering process, participants were asked to respond to a questionnaire and then use the software. The results of questionnaire for the academic group including 55 students at London Metropolitan University is shown in figure 3.

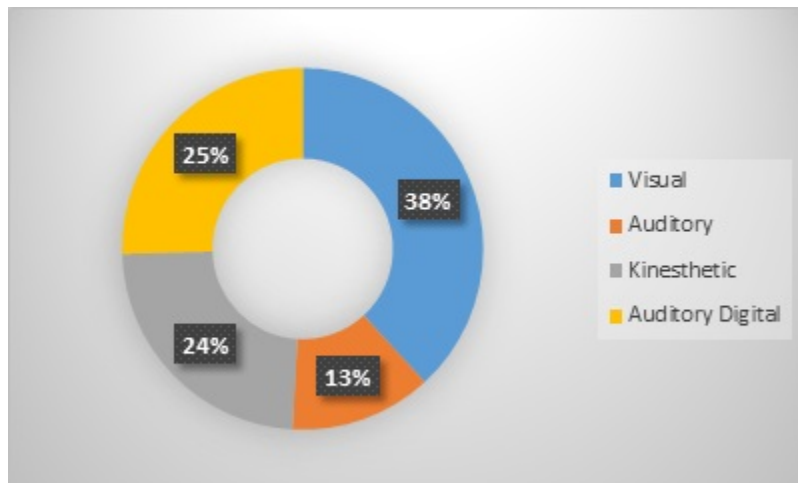


Fig. 3 Questionnaire results for the preferred representational system of participants

Analysing the questionnaire results, it was noted that the preferred representational system for 21 participants was visual, for 7 participants auditory, for 13 participants kinesthetic and for 14 participants auditory digital. According to figure 3, 38 percent of participants were visual, 13 percent were auditory, 24 percent were kinesthetic and 25 percent were auditory digital. The achieved results by the software for the academic group of participants is shown in figure 4.

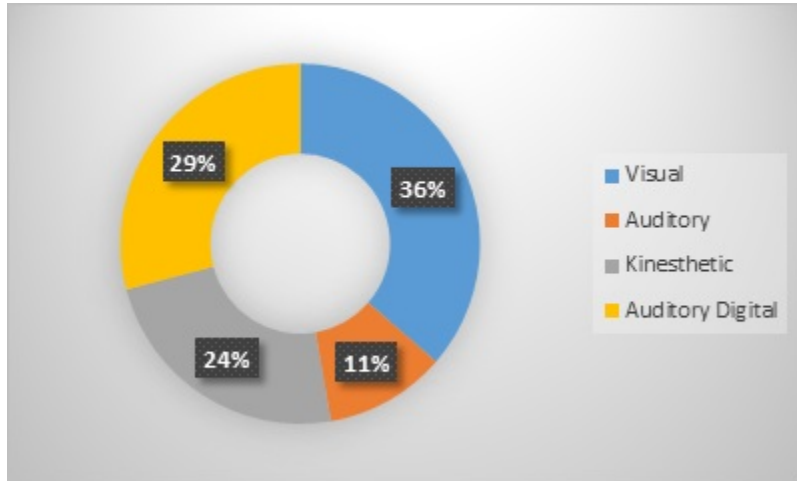


Fig. 4 Software results for the preferred representational system of participants

After analysing the software results, 20 participants have been identified with visual preference, 6 participants with auditory preference, 13 participants with kinesthetic preference and 16 participants with auditor digital preference. According to figure 4, 36 percent were visual, 11 percent auditory, 24 percent kinesthetic and 29 percent auditory digital. As a result, we have compared the result of manual questionnaire with the obtained result by the software and this is shown in figure 5.

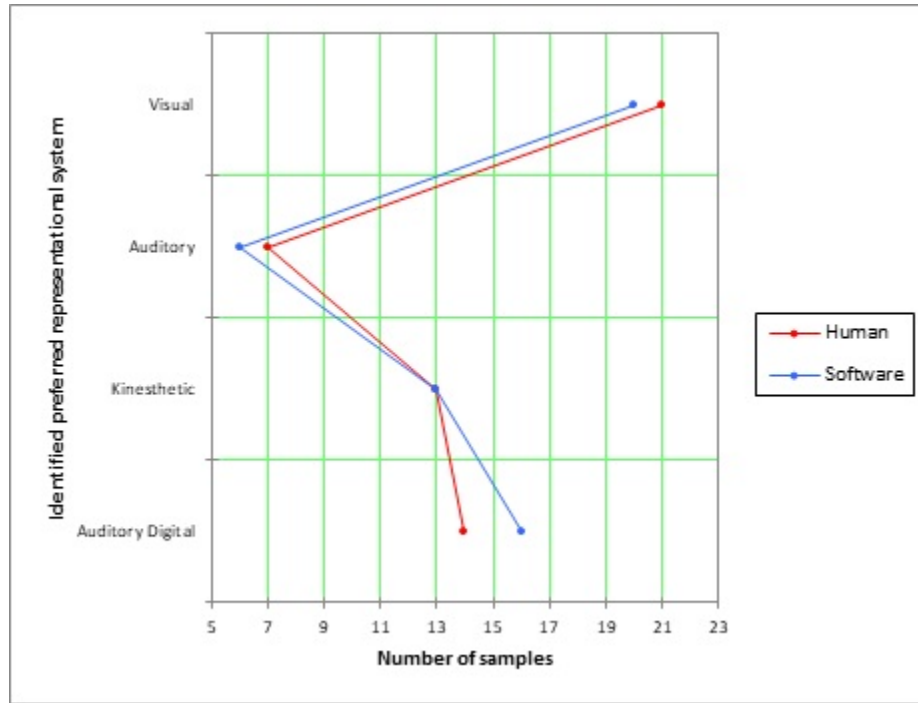


Fig. 5 Comparing the number of preferred representational systems identified by a questionnaire and software.

As shown in figure 5, the numbers of participants with a similar preferred representational system is very close to each other and the result of manual questionnaires are very similar to the obtained results by the software with only one subject being in different categories of representational systems comparing the manual questionnaire with the software. Figure 5 shows that the questionnaires' results identify 21 participants with preference of the visual representational system while the software's results identify 20 participants as visual people. Figure 5 also shows that the number of kinesthetic participants is the same for the questionnaire and the software. Furthermore, there is a difference of 2 people between the questionnaire's results and the software's results for auditory digital participants.

According to table 5, the manual questionnaire has recognised 38.18 percent of participants with visual preference, 12.72 percent with auditory preference, 23.63 percent with kinesthetic

preference and 25.45 percent with auditory digital preference. While the software has recognised 36.36 percent with visual preference, 10.90 percent with auditory preference, 23.63 percent with kinesthetic and 29.09 percent with auditory digital preference. This reveals 1.82 percent difference between the manual questionnaire and the software results for the visual preference and 1.82 percent difference for auditory preference. For kinesthetic preference, the percentage for the software is exactly the same as manual questionnaire. Finally, the difference between the software and the manual questionnaire for auditory digital preference is 3.64 percent. These percentages could provide further proof that the software performs well.

Table 5 Difference between questionnaire and software results

	Visual	Auditory	Kinesthetic	Auditory Digital
Questionnaire	38.18%	12.72%	23.63%	25.45%
Software	36.36%	10.90%	23.63%	29.09%
Difference	1.82%	1.82%	0%	3.64%

3.2 Comparing the results of software with human NLP practitioner

The number of predicates or words that have been referring to each one of the representational systems in the participants' answers were considered. There were 974 words identified as a predicates in the questionnaire responses and the result for each on of representational systems is shown in figure 6.

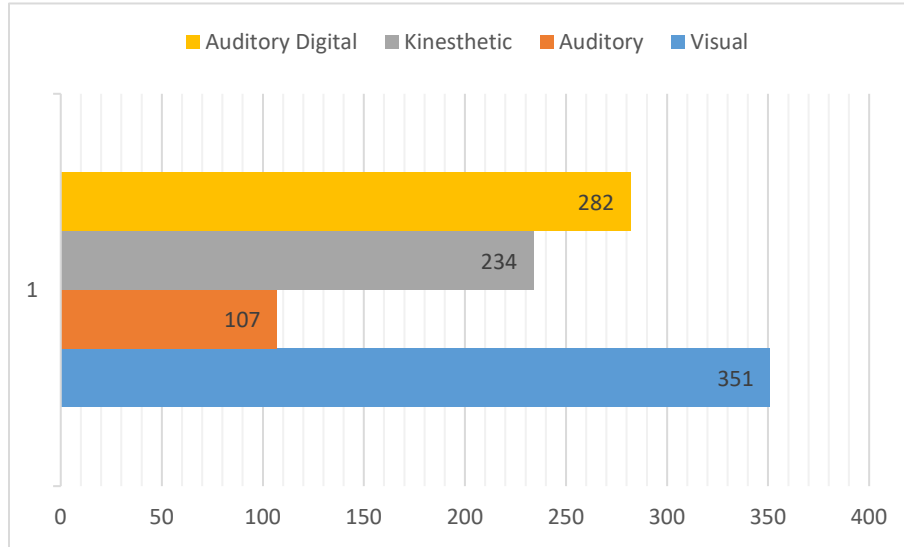


Fig. 6 Number of predicates related to each representational system identified by human

Results show that out of 974 predicates, 351 predicates were related to the visual representational system, 107 predicates were related to auditory, 234 predicates related to kinesthetic and 282 predicates were related to the auditory digital representational system. Table 6 shows that 36.03 percent of words have been related to visual preference, 10.98 percent related to auditory, 24.02 percent related to kinesthetic and 28.95 percent related to auditory digital.

Table 6 Recognised predicates by human

	Visual	Auditory	Kinesthetic	Auditory Digital
Number of words	351	107	234	282
Percentage	36.03%	10.98%	24.02%	28.95%

The number of predicates identified by the software were 978 words. The results are shown in figure 7.

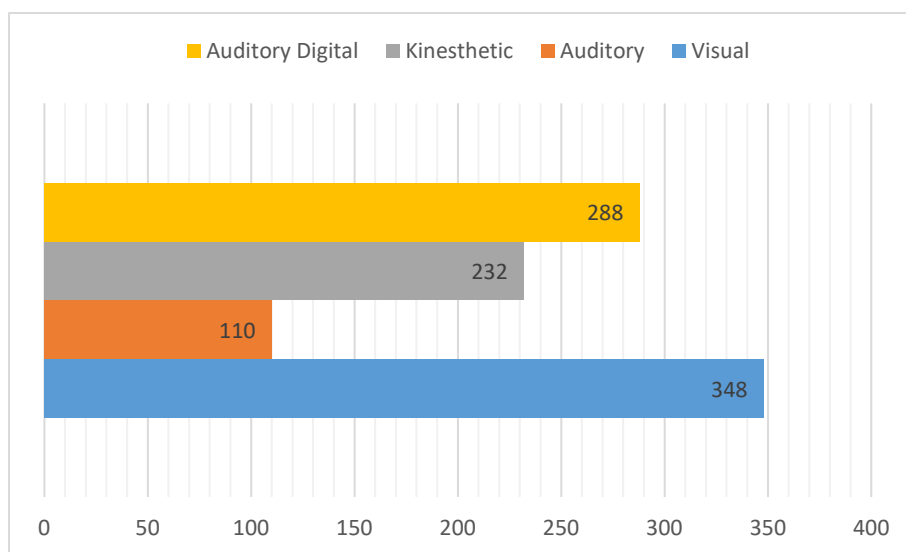


Fig. 7 Number of predicates related to each representational system identified by software

As presented in figure 7, out of 978 predicates, 348 were related to the visual representational system, 110 predicates were related to auditory, 232 were related to kinesthetic and 288 predicates were related to the auditory digital representational system. Table 7 shows that 35.58 percent of words have been related to visual preference, 11.24 percent related to auditory, 23.72 percent related to kinesthetic and 29.44 percent related to auditory digital.

Table 7 Recognised predicates by software				
	Visual	Auditory	Kinesthetic	Auditory Digital
Number of words	348	110	232	288
Percentage	35.58%	11.24%	23.72%	29.44%

Finally, the number of predicates identified manually were compared to the number of predicates identified by software and the results shown in figure 8.

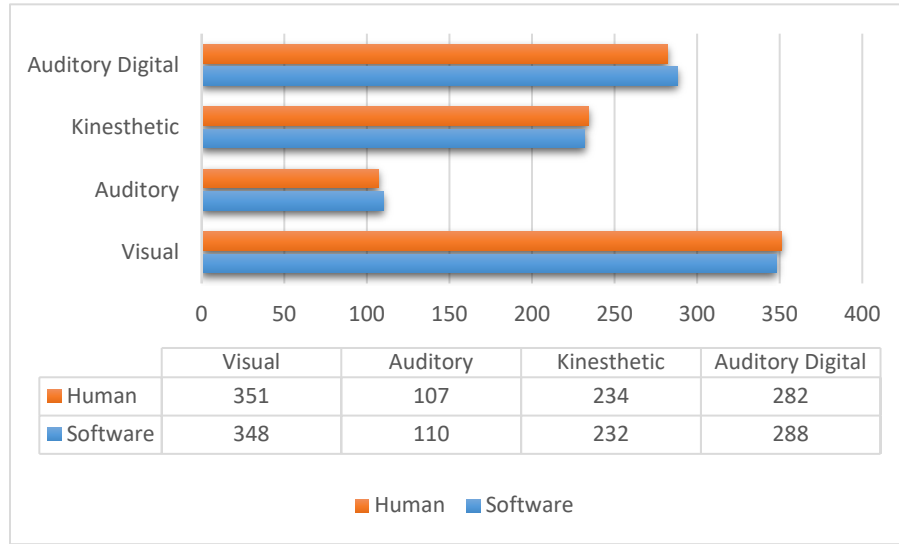


Fig. 8 Comparing the number of predicates identified by human and software

The most notable difference in the increased number of predicates recognised by the software in comparison to the manual questionnaire conveying increased accuracy and competence of the intelligent software. According to figure 8, for the visual representational system, the manual has been relatively better than the software in terms of identifying the relevant predicates where 351 predicates have been identified manually and 348 predicates identified by the software. For the auditory representational system, 107 predicates have been identified manually and 110 predicates were identified by the software showing better performance by the latter. For the kinesthetic representational system, the data is almost similar to the visual representational system and the manual system has been slightly better than the software whereby 234 predicates have been identified manually while 232 predicates have been identified by the software. Finally, for the auditory digital representational system, the software has been more successful identifying 6 more predicates. The number of identified predicates by the questionnaire was 282 while the software recognised 288. Table 8 shows the percentages and the difference between the manual performance and the software.

Table 8 Comparing the human and software performance

	Visual	Auditory	Kinesthetic	Auditory Digital
Human	36.03%	10.98%	23.63%	28.95%
Software	35.58%	11.24%	23.63%	29.44%
Difference	0.45%	0.26%	0%	0.49%

Moreover, the number of visual predicates for each person, recognised by software and human were calculated separately and a ‘time series visualisation’ diagram has been used to compare the results. This is shown in Figure 9.

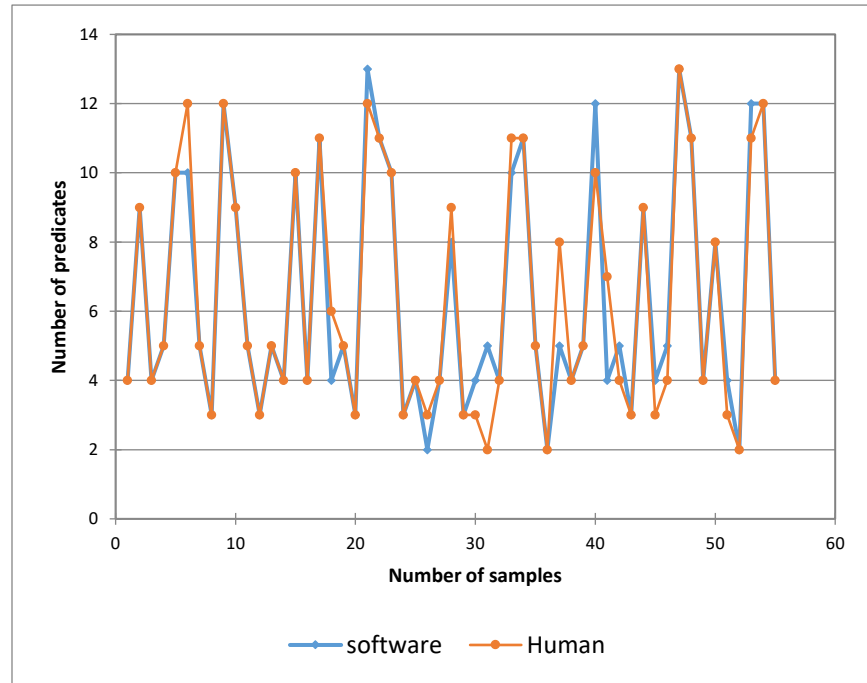
**Fig. 9** Comparing the number of visual predicates, identified by the human and software

Figure 9 shows 39 out of 55 people have been identified with the same number of visual predicates. For 10 cases, the difference in the number of recognised visual predicates was only 1 word. This difference for 3 of the cases was 2 words, and for the remaining three, by 3 words.

As mentioned before, figure 5 shows that out of 55 people, the software identifies 20 people with the visual preferred representational system while results acquired by humans recognised 21 cases. After considering all those people who have been identified with the visual preference and also according to figure 14, it has been realized that both human and software have identified the visual preference for the same cases, even if there is a small difference between the number of recognised visual predicates for each person.

Figure 10 shows the result of comparing the number of recognised auditory predicates for each person, identified by software and human.

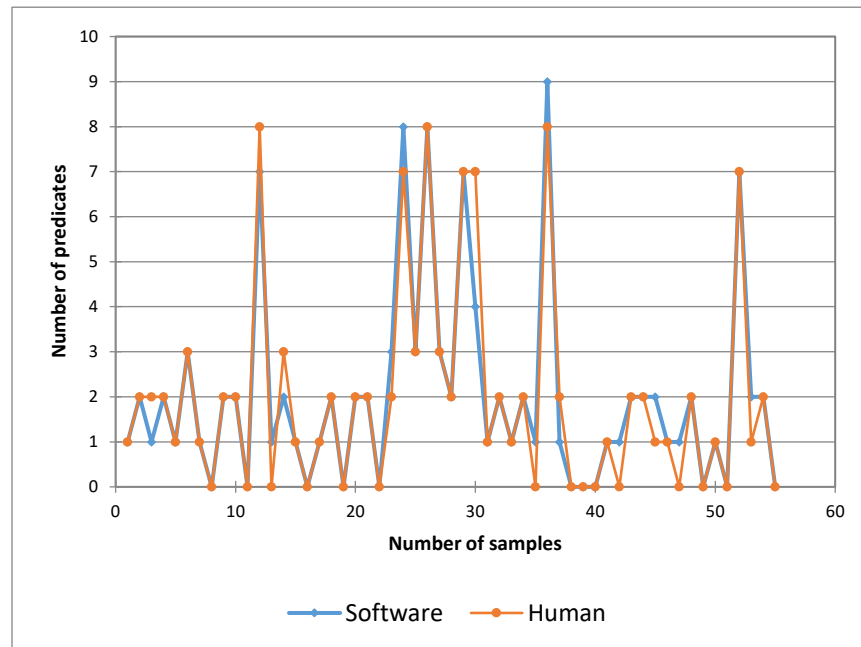


Fig. 10 Comparing the number of auditory predicates, identified by the human and software

Figure 10 shows that 41 out of 55 people have been identified with the same number of the recognised auditory predicates. 13 people have been recognised with only one word difference and only one person has been identified with 3 words difference.

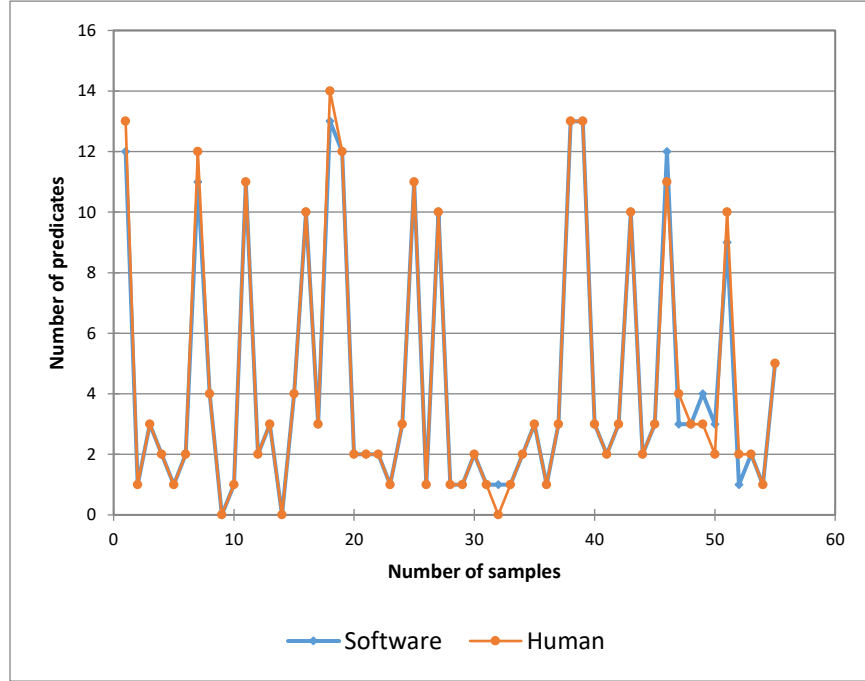


Fig. 11 Comparing the number of kinesthetic predicates, identified by the human and software

According to figure 11, the performance of software and human in regards to identifying the kinesthetic preference has been very similar and the number of people who have been identified with the same number of recognised kinesthetic predicates is 45 out of 55. Figure 11 shows that the other 10 people have been identified with only one word difference.

Finally, figure 12 shows that 35 people out of 55 people have been identified with the same number of recognised auditory digital predicates and there are 17 people who have been identified with only one word difference between the software results and human results. One person has been identified with 2 words difference, one person with 3 words difference and one person with 4 words difference.

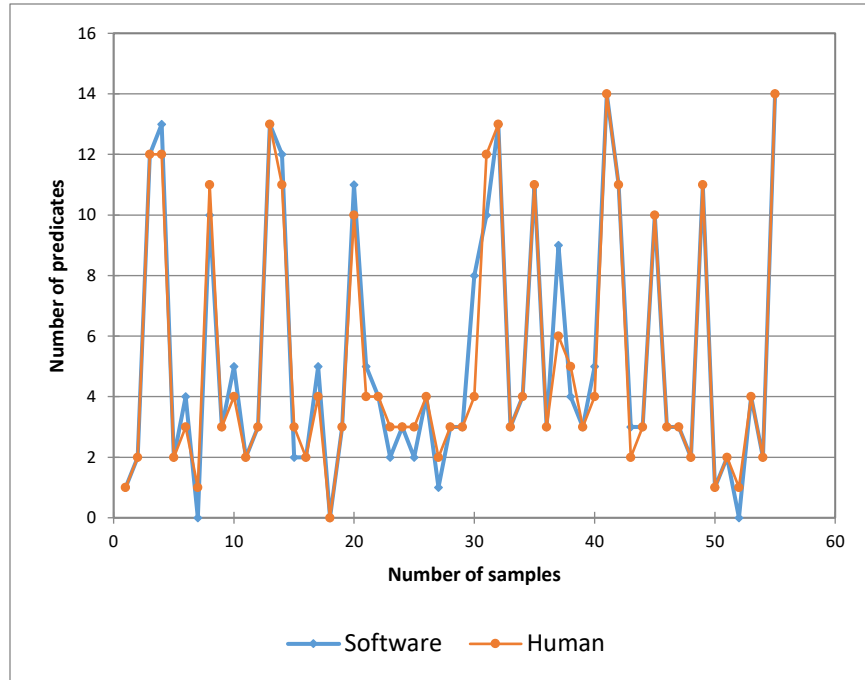


Fig. 12 Comparing the number of auditory digital predicates, identified by the human and software

After analysing all predicated recognised by both software and human, three main factors have been recognised to contribute to the difference between the number of predicates related to each representational system, identified by the software and human. The first reason is the limitation of words in the software dictionary, which impacted efficiency of the software. The second factor contributing to the difference of performances is the human errors involved leading to some predicates being identified by the software but not by humans. Finally, synonyms of words used in some cases had been recorded in association to other representational systems. As a result, the recognised predicate has been considered as an indicator for one of the other representational systems and not for the one which has been identified by human.

Overall results suggest that the software is able to replicate the human performance and recognise the preferred representational system correctly based on the language used by an individual. The results also present more accuracy in performance in some aspects of this process

in comparison to the manual alternative.

However, there was an important issue to be considered during the data gathering process by the software. The efficiency of the automatic software increases by a higher word count of answers where the response includes more detail, and consequently more words. Participants were informed about this and asked to respond to each question clearly explaining their answers vividly. Additionally, the questions were designed in a way to encourage participants to give full answers, using the maximum number of words. Nonetheless, some of our participants were responding to some of questions in short sentences with the minimum words. This was expected to cause an issue.

Interestingly, despite some participants responding to some of the questions with very few words, the overall results were accurate and corresponded to the results of the manual questionnaire. This could be due to the number of questions and also the purpose of each question. Moreover, the questions were considerate of various dimensions of personality, with each question assessing the participant from a different aspect. Thus, the individual must think about one concept from different perspectives. As a result, accuracy was achieved through the combining of responses, even if some answers were short.

3.3 Comparing the results of software and questionnaire for the industrial group

The software was used to identify the preferred representational system for each employee at UKSTUDY Company. They were also asked to respond the explained questionnaire in section 2.2. Figure 13 shows the number of employees in each category identified by the software and the questionnaire.

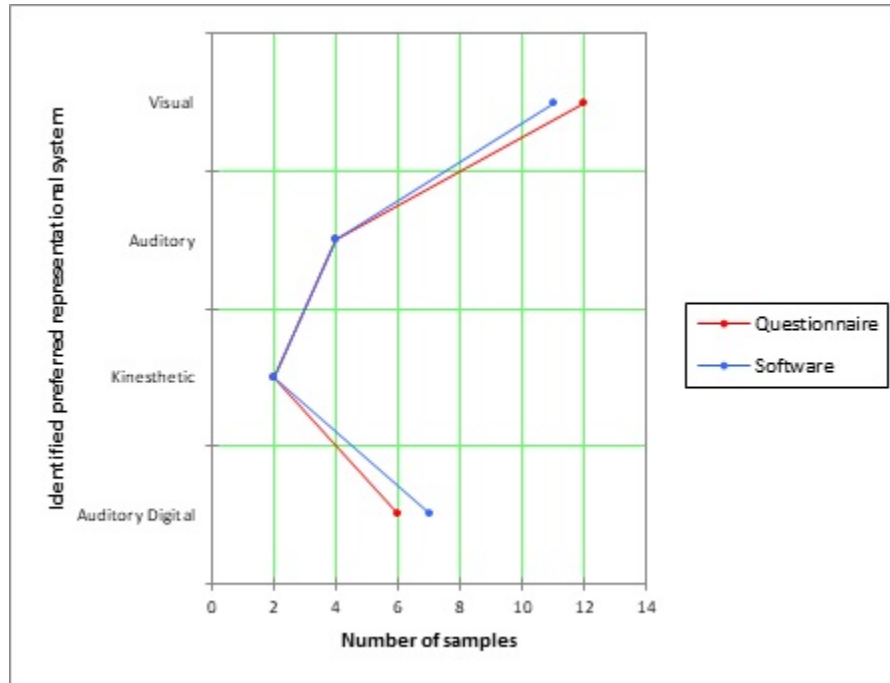


Fig. 13 Number of employees in each category of preferred representational system

After analysing the software results for the 24 employees in the company, 11 employees were identified with visual representational system as their preferred representational system, 4 employees were identified with auditory, 2 employees with kinaesthetic and 7 employees with auditory digital representational system. On the other hand, analyzing the questionnaire results, it was noted that the preferred representational system for 12 participants was visual, for 4 participants auditory, for 2 participants kinesthetic and for 6 participants auditory digital. This outstanding results shows that the performance of the software in identifying the auditory and kinaesthetic representational system was exactly the same and there was only one person difference for visual and auditory digital representational systems.

Table 9 shows the comparison between the percentage of each preferred representational system in the company identified by the software and the questionnaire.

Table 9 Difference between questionnaire and software results

	Visual	Auditory	Kinesthetic	Auditory Digital
Questionnaire	50%	17%	8%	25%
Software	46%	17%	8%	29%
Difference	4%	0%	0%	4%

According to table 9, the software has recognised 46 percent of participants with visual reference, 17 percent with auditory reference, 8 percent with kinaesthetic preference and 29 percent with auditory digital preference. While the questionnaire has recognised 50 percent with visual preference, 17 percent with auditory preference, 8 percent with kinaesthetic preference and 25 percent with auditory digital preference. This reveals 4 percent difference between the software and the questionnaire results for the visual and auditory digital preference. For auditory and kinesthetic preference, the percentage for the software is exactly the same as manual questionnaire. These percentages could provide further proof that the software performs well.

According to Figure 13 visual representational system was the most popular representational system in the company and the second popular representational system was auditory digital. Figure 13 also shows that third preferred representational system for the employees was auditory and the fourth one was kinaesthetic which had the lowest popularity in the company. The software successfully provided the relevant solutions to each employee, in order to improve their communications with their colleagues. The software also provided the relevant solutions for visual representational system as the most popular representational system, in order to improve the communication through the company and improve the staff learning and understanding level in the meetings.

3.4 Comparing the software with previous attempts for automation

The performance of the software was also compared with some of the previous attempts. As it was discussed in section 1, the previous attempts to automate the identification process of the preferred representational system have only been as competent as a simple computerization of the concept rather than intelligent automation. For instance, One of the few efforts for NLP automation is an online NLP test on <http://www.vaknlp.com>. This website makes an effort to explain the relationship between the human senses and different types of personality. This test focuses on representational systems including the visual, auditory and kinesthetic types. Figure 14 shows that a collection of 10 questions is used to identify the preferred representational system of the user.

1. Your main interests:
☐ You know fine wines.
☐ You are talented working with your hands.
☐ You prefer taking souvenir photos.

2. Your sense of humor:
☐ Sarcastic.
☐ Subtle.
☐ Sometimes mocking.

3. The image you want to project:
☐ Well spoken and trustworthy.
☐ Sane with balanced interests.
☐ Wise from life's experiences.

4. When you don't want to do anything:
☐ You listen to a CD or to the radio.
☐ You watch television or read magazines.
☐ You take a long, long bath.

5. In general with pets:
☐ You watch each other carefully.
☐ You pet them affectionately.
☐ You play with them while talking to them.

6. Your negative characteristics:
☐ To be a pushover.
☐ To be selfish.
☐ To be rigid.

7. Your influence:
☐ You are persuasive.
☐ You engage people's feelings.
☐ You present an attractive image.

8. To make you a new friend:
☐ You are sweet and charming.
☐ You are funny and make them laugh.
☐ You compliment them and are solicitous.

9. For you, a good film must:
☐ Make you dream.
☐ Be intelligent.
☐ Be a beautiful spectacle.

10. Your friends like you because:
☐ You entertain them.
☐ You help them to relax.
☐ They have great conversations with you.

Calculate the results Clear all

Fig. 14 VAK test (<http://www.vaknlp.com>, 2016)

During this research, these 10 questions were tested by different users at different times. However, it was recognised that this survey is very limited in accuracy to identify the profile of

the respondents. An example of this is shown in figure 15.

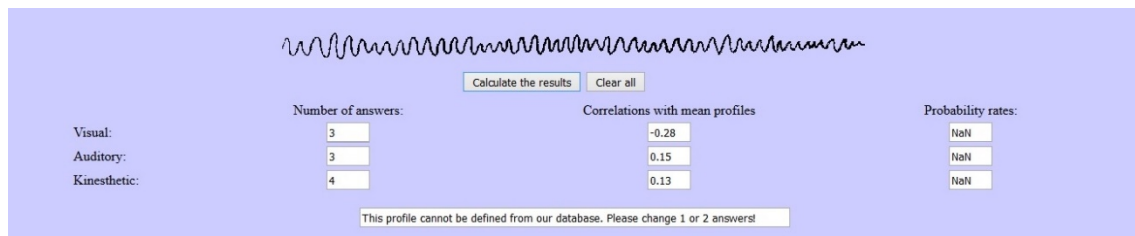


Fig. 15 VAK test result (<http://www.vaknlp.com>, 2016)

It was recognised that this software is able to identify how many answers were related to each one of visual, auditory and kinesthetic representational systems. The user's preference is then identified based on the highest number of answers correlated with a representational system, based on associations of answers and systems previously defined for the software. Thus, it seems that for each set of numbers, a profile has been defined for the software. Its function would therefore, stem from comparing the acquired set of numbers with the defined set of numbers and followed by the display of the result as the user's preference. However, as shown in figure 15, if the pattern of numbers acquired from the user is different to the defined set of numbers, the system is unable to identify the user's preference. Subsequently, the system would ask the user to change some of their answers. This results in a significant impact on the accuracy and reliability of the results obtained. Another important issue is that this website does not consider auditory digital representational system.

Another test named 'NLP Representational System Preference Test' was also analyzed in this research. This test is available online on the 'Transform Destiny' company's website. This company is working in the field of NLP Training, organizing various events and running online courses on NLP. Figure 16 shows the format of this test.

https://www.transformdestiny.com/nlp-guide/nlp-representational-systems-preference-test.asp

4 = **Closest** to describing you
 3 = Next best description
 2 = Next least descriptive
 1 = **Least** descriptive of you

1. I make important decisions based on: ★ 1

☐ gut level feelings
☐ which way sounds the best
☐ what looks best to me
☐ precise review and study of the issues

2. During an argument, I am most likely to be influenced by:

☐ the other person's tone of voice
☐ whether or not I can see the other person's point of view
☐ the logic of the other person's argument
☐ whether or not I am in touch with the other person's true feelings

3. I most easily communicate what is going on with me by:

☐ the way I dress and look
☐ the feelings I share
☐ the words I choose
☐ my tone of voice

4. It is easiest for me to:

☐ find the ideal volume and tuning on a stereo system
☐ select the most intellectually relevant point in an interesting subject
☐ select the most comfortable furniture
☐ select rich, attractive color combinations

5. Select the most accurate:

☐ I am very attuned to the sounds of my surroundings
☐ I am very adept at making sense of new facts and data

Fig. 16 NLP Representational System Preference Test (Transform Destiny Company, 2016)

This was a very simple test including 5 questions. Each question provided four answer options requiring rating of these options by choosing a number from 1 (As the least descriptive of you) to 4 (As the closest one to describe you). The software presents the result in a table (Figure 17).

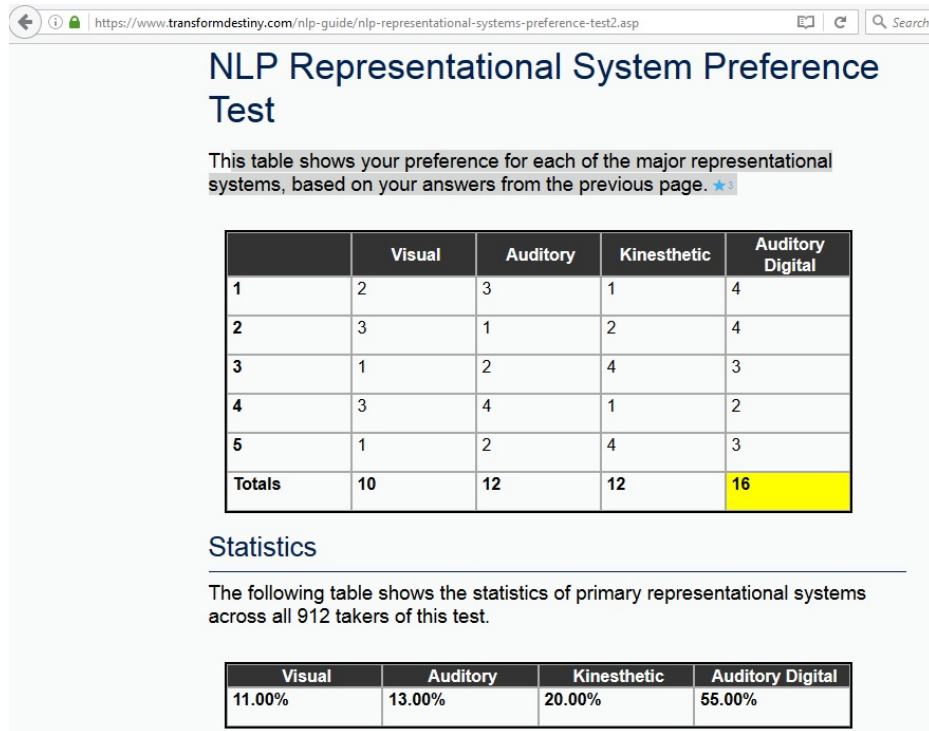


Fig. 17 NLP Representational System Preference Test result (Transform Destiny Company, 2016)

The table of results consists of the test taker's response to each question and the total score for each category. The highest total score would be indicative of the preferred representational system of the test taker. The software also presents the statistics of primary representational systems across all previous respondents of this test. In overview, it can be said that the structure of this test is basic in considering 5 questions for analysis of character. Moreover, the test provides discrete options to be chosen from rather than allowing a more candid expression in response to questions. It can also be argued that the methodology of this test lacks intelligence. Similarly to online self-assessments, results are dependent on the individual's judgment and opinion of themselves.

On the other hand, the software developed in this research gives an opportunity to the participants for self-expression, instead of limiting them with pre-set discrete options to be chosen

from. The software is also able to use an intelligent method in order to analyse and judge on categorization of the participants' response without relying on the judgement of participants.

Another shortcoming of the available online surveys addressed is the possibility of establishing a conversation with a human being, or in other words, the ability of computer-human interaction. While creating a more user-friendly platform in allowing for personalised response to questions, this computer-human interaction also allows for development of answers throughout an interactive conversation. Meanwhile, the software is able to direct the conversation towards clarifying ambiguous or unspecific answers towards ensuring an accurate judgement on categorizing the individual's preference of representational system using Natural Language Processing techniques.

Another limitation of previous methods is the consideration and analysis for the visual, auditory and kinesthetic representational systems only, leaving out a relatively popular category of the auditory digital representational system. The developed software in this research, however, examines and explores for all four general representational systems to recognize the preferred representational system. Moreover, the previous attempts of NLP assessments are limited to identifying the preferred category of representational systems alone, whereas the developed software in this research provides a series of solutions to improve the efficacy of communication and learning methodologies of the individual, based on their identified preferred representational system.

Another software named 'Manteya Email' was also analysed which utilises the NLP representational systems theory. This software was created by Manteya, the online persuasion experts, and it is also available to reach on the ANLP (The Association for Neuro Linguistic Programming) website. This software is capable of analysing the model of writing used by

individuals in their emails in order to aid them to communicate more effectively. A powerful psychological database is used for this process to understand how people build rapport through computers when they are communicating via emails (ANLP, 2016). Neil Trigger who is the founder of Manteya, claims that he utilised a unique PhD research for this application where he has found a method for analysing incoming Emails and scoring their content according to psychological methods. According to the ANLP website (2016) “When you reply to an email, it automatically cross-references your email with the score the original sender has accumulated over time. If there is a mismatch identified, the system will inform you of what you need to change to make it more persuasive.” Hence, the software can improve interactions with anyone you communicate with in trying to improve the psychological impact of your email.

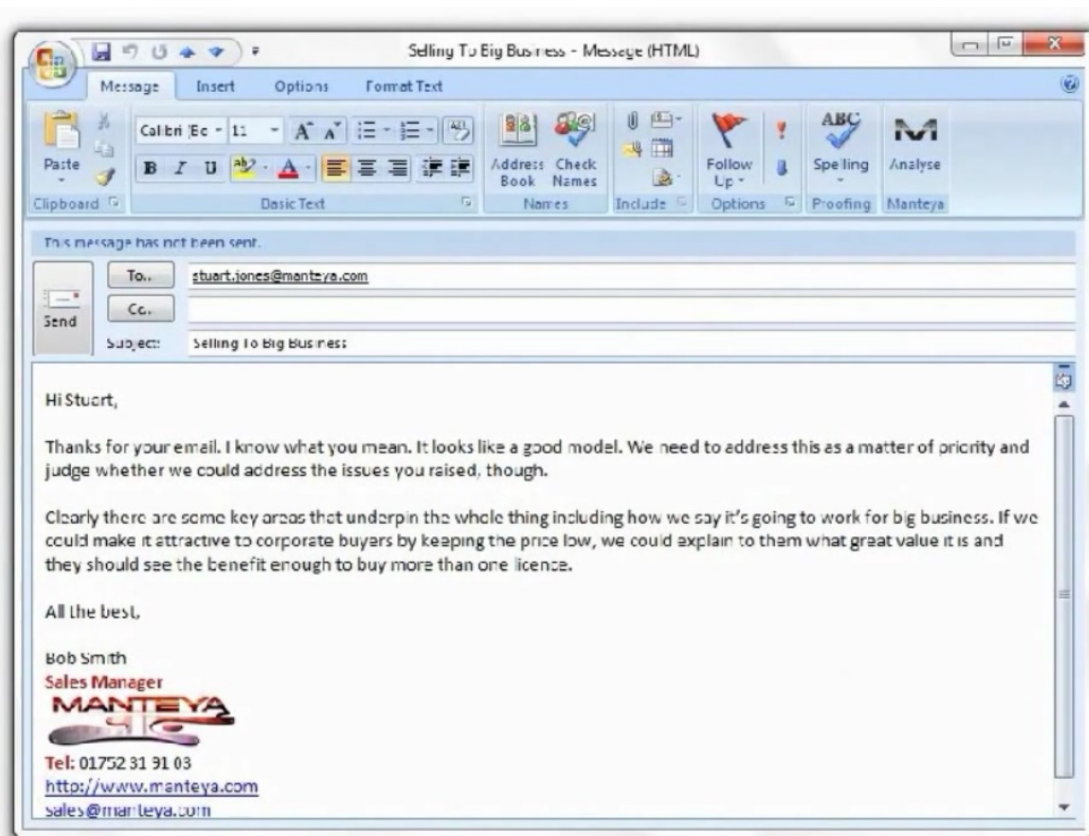


Figure 18: A sample of using Manteya in outlook

Figure 18 is a screen shot of a video on YouTube that describes how the Manteya email software works. In this email, the software would recognise specific words such as ‘see’, ‘clearly’, ‘explain’, ‘looks’, ‘say’ and etc. in the text. Then, these words will be analysed and upon clicking the Manteya button, a report about the text would be displayed as shown in Figure 19.

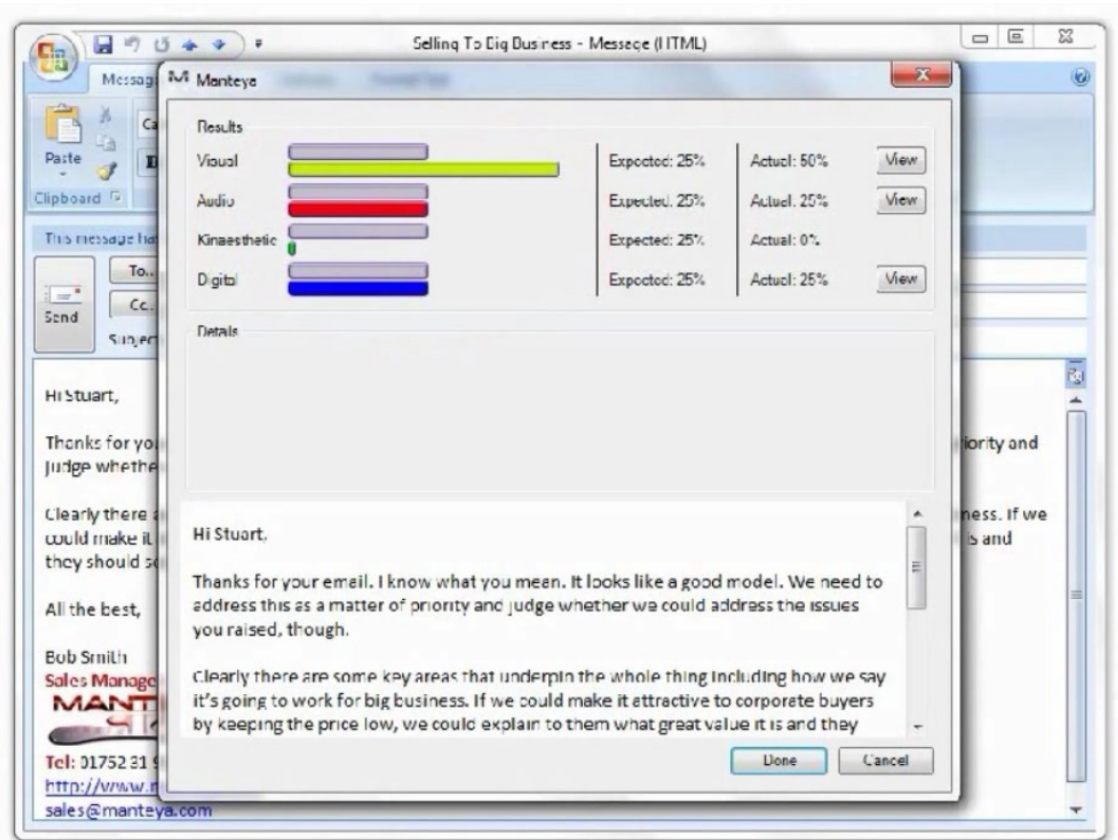


Figure 19: Manteya result

Figure 19 shows that the written text has more emphasis on words that are associated to use by visual people. By clicking on the view button, the software would show how you may change the text to make it more psychologically effective. Then, by clicking on the update button, the software would change the text automatically replacing words with the suggested words to improve elements of the language. The result is shown in figure 20.

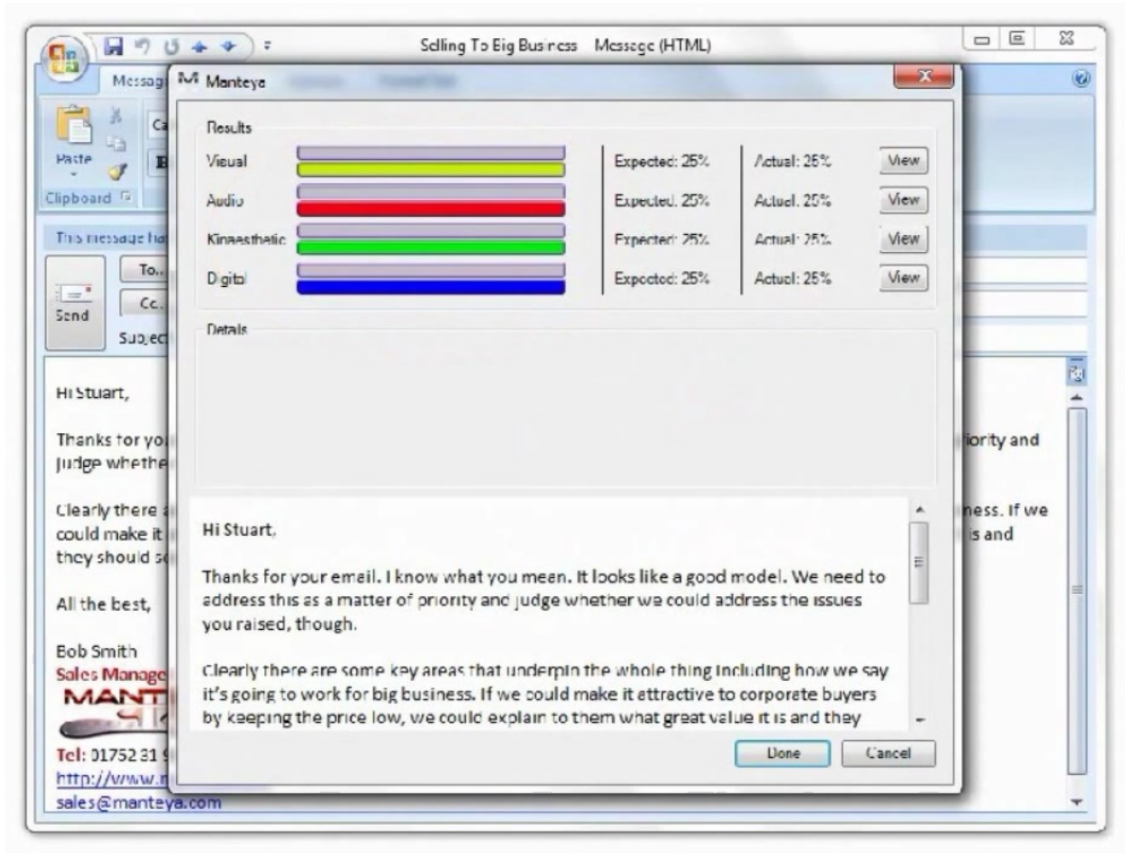


Figure 20: Manteya result

Figure 20 shows the result of this change in improvement of language in using a wider range of vocabulary associated to all representational systems. A concern in this correction is the standardisation of the language to one type where all representational systems are embodied equally. This balance is, however, not considerate of the recipient of the email where a person may be extremely auditory and not so responsive to other representational systems. In reality every individual has a different balance of preferences with varying levels of understanding through each representational system, making the process of standardisation to one specific balance ineffective for its objective. Another shortcoming of this software is the limitations of the dictionary used by the software which was found to be incapable of recognising many predicates, especially those associated to the auditory digital representational system. The proposed methodology in this

research has overcome this limitation with expanding the dictionary used, recognising relevant synonyms of predicates of each category of representational system. As a result, the software is able to perceive the language used by the user better.

The limitations of previous attempts discussed would imperatively impact the accuracy and reliability of the results obtained. Comparisons discussed between the developed software in this research and previous attempts for automation, outlines various strategies implemented in the developed software to address and improve on the limitations of previously available assessments. Hence the accuracy and reliability of the software is strategically upgraded and better than the previous attempts.

4. Conclusion

This research has automated the process of identifying the preferred representational system which is one of the most important aspects of Neuro Linguistic Programming and vital during the personality development process. Natural Language Processing, a subfield of artificial intelligent was used as a tool for the automation process. As a result, a new intelligent software which is able to act like an experienced psychologist or NLP practitioner has been developed based on Python and Natural Language Processing Tool-Kit (NLTK). The Software has been tested on an academic group of participants including 55 students at London Metropolitan University who have been studying different subjects at PhD research, undergraduate and postgraduate levels. It has also been tested on an industrial group of participants including 24 employees in UKSTUDY Company located in Brighton. They were asked to respond to a manual questionnaire, designed in order to understand their preferred representational system. The obtained results by the software were

compared with the attained results by the questionnaire, demonstrating a superior performance and a high level of accuracy and reliability of the software against the manual questionnaire. Furthermore, in recognizing the language, recognising the predicates related to each representational system, and identifying the preferred representational system, the performance of the software was compared with a human NLP practitioner and it was shown to be slightly more accurate than the results from the NLP practitioner. Moreover, the performance of the software was compared with previous automation attempts and it was shown that the proposed methodology and the developed software in this research has overcome the significant limitations of previous attempts. Based on the results, it can be concluded that the proposed software is more robust in identifying auditory digital and auditory representational systems than an experienced NLP practitioner. In contrast, it could also be concluded that the developed software is slightly less effective for visual and kinesthetic representational system analysis compared to a human NLP practitioner. Therefore, the novel methodology presented in this research could successfully improve the accuracy and reliability of the identification process for the preferred representational system with the advantage of significantly decreasing the inaccuracies associated with the manual processes such as, the lack of experience, personal judgment, different level of skills and other human errors.

5. Compliance with Ethical Standards

Conflict of Interest: The authors declare that they have no conflict of interest.

Ethical approval: For this type of study, formal consent is not required.

References

- ANLP (2016) Improving your communication using Manteya Email. <https://anlp.org/business-resources-for-nlp-professionals/manteya-email>. Accessed 18 Dec 2017
- Bensted C (2014) Representational systems. <http://badis.co.uk/resources/Repsys.pdf>. Accessed 12 Feb 2018
- Bird S, Klein E, Loper E (2009) Natural Language Processing with python: analyzing text with the natural language toolkit. O'Reilly Media Inc, Newton
- Brefi Group Limited (2004) Representational system. http://www.brefigroup.co.uk/acrobat/nlp_representational_systems.pdf. Accessed 24 Feb 2018
- Casale P (2012) NLP secrets: upgrade your mind with Neuro Linguistic Programming. Creative Media NZ Ltd, Auckland
- Chopra A, Prashar A, Sain A (2013) Natural Language Processing. Int J Technol Enhanc Emerg Eng Res 1(4):131–134
- Chowdhury G (2003) Natural Language Processing. Ann Rev Inf Sci Technol 37(1):51–89
- Dale R (2010) Classical approaches to Natural Language Processing, handbook of Natural Language Processing. CRC Press, Boca Raton
- Ellerton R (2007) NLP's Auditory digital representational systems. <http://www.renewal.ca/nlp48.html>. Accessed 27 Jan 2018

Ellerton R (2015) Modalities and representational systems. <http://www.renewal.ca/nlp10.htm>.

Accessed 3 Feb 2018

Gudwin R (2000) Evaluating intelligence: a computational semiotics perspective. In: IEEE international conference on systems, man and cybernetics, Nashville, Tennessee, USA

Hervás R, Bautista S, Rodríguez M, Salas T, Vargas A, Gervás P (2012) Integration of lexical and syntactic simplification capabilities in a text editor. *Procedia Comput Sci* 27:94–103

Horn D (2008) Syntactic structures in language and biology. *Cogn Process* 9(14):153–158

Juiced Concepts Limited (2012) NLP home study programme (V2.0). <https://s3-eu-west-1.amazonaws.com/nlpraconlinemp3/RepSystems/RepresentationalSystems.pdf>. Accessed 5

Mar 2018

Jurafsky D, Martin J (2014) *Speech and Language processing*. Prentice Hall, Pearson Education International, Upper Saddle River

Lazarus J (2010) *Successful NLP: for the results you want*. Crimson Publishing, Surrey

Linder-Pelz S (2010) ‘NLP coaching’: an evidence-based approach for coaches, leaders and individuals. Koganpage, London

Manning CD, Raghavan P, Schütze H (2009) *Introduction to information retrieval*. Cambridge University Press, Cambridge

McAfee K (2014) ‘The power of words’, an introduction to NLP representational systems, becoming a more effective communicator. <http://www.americasmarketingmotivator.com/wp->

[content/uploads/2014/09/Workbook-The-Power-of-Words-NLP-Representational-Systems.pdf](http://www.distancelearning.academy/wp-content/uploads/2014/09/Workbook-The-Power-of-Words-NLP-Representational-Systems.pdf). Accessed 19 Dec 2017

Monkeypuzzle Training and Consultancy (2016) 'The power of your senses'. Using NLP representational systems to improve how you communicate, relate and learn. <http://www.monkeypuzzlettraining.co.uk/free-documents/The-Power-of-Your-Senses-NLP.pdf>. Accessed 21 Jan 2018

NLP Dynamics Ltd. (2013) Representational systems. <http://www.distancelearning.academy/wp-content/uploads/2015/02/Representational-Systems.pdf>. Accessed 16 Feb 2018

NLTK official website (2016) NLTK 3.0 documentation. <http://www.nltk.org/#natural-language-toolkit>. Accessed 4 Mar 2018

O'Connor J, Seymour J (1993) Introducing Neuro Linguistic Programming: psychological skills for understanding and influencing people, Revised edn. The Aquarian Press, London

Palmiero M, Di Matteo R, Belardinelli M (2014) The representation of conceptual knowledge: visual, auditory, and olfactory imagery compared with semantic processing. *Cogn Process* 15(2):143–157

Rayner Institute (2015) The representational systems. http://www.raynerinstitute.com/uploads/9/8/6/1/9861170/nlp_rep_system.pdf. Accessed 16 Feb 2018

Singer MT, Lalich J (1996) 'Crazy therapies'. What are they? Do they work?. Jossey-Bass, San Francisco

Tosey P, Mathison J (2003) Neuro-linguistic programming and learning theory: a response. Curric J 14:371–388

Tosey P, Mathison J (2006) Introducing Neuro-Linguistic Programming Centre for management learning & development. School of Management, University of Surrey, Guildford

Transform Destiny (2015) Introduction to neuro-linguistic programming. <http://www.freenlphomestudy.com/membersonly/iNLP/iNLPMannual.pdf>. Accessed 15 Jan 2018

Transform Destiny (2016) NLP Representational System Preference Test. <https://www.transformdestiny.com/nlp-guide/nlp-representational-systems-preference-test.asp>. Accessed 2 Feb 2018

Webster J, Kit C (1992) Tokenization as the initial phase in NLP. In: The 15th international conference on computational Linguistics, France

Witkowski T (2010) Thirty-five years of research on neuro-linguistic programming. NLP research data base. State of the art or pseudoscientific decoration? Pol Psychol Bull 41(2):58–66